

From algorithmic computing to direct retrieval: Evidence from number and alphabetic arithmetic in children and adults

PIERRE BARROUILLET and MICHEL FAYOL
Université de Bourgogne, Dijon, France

A number of theories of mental arithmetic suggest that the ability to solve simple addition and subtraction problems develops from an algorithmic strategy toward a strategy based on the direct retrieval of the result from memory. In the experiment presented here, 2nd and 12th graders were asked to solve two tasks of number and alphabet arithmetic. The subjects transformed series of 1 to 4 numbers or letters (item span) by adding or subtracting an operand varying from 1 to 4 (operation span). Although both the item and operation span were associated with major and identical effects in the case of both numbers and letters at 2nd grade, such effects were clearly observable only in the case of letters for the adult subjects. This suggests the use of an algorithmic strategy for both types of material in the case of the children and for the letters only in the case of the adults, who retrieved numerical results directly from memory.

During the last 20 years, a considerable body of research has been devoted to the study of the development in processing of simple addition problems (e.g., $3+2$) (Ashcraft, 1982, 1992). This research has shown that adults possess well-developed mental representations of numerical facts which enable them to retrieve the results from memory (e.g., $3+2=5$) instead of having to calculate them (Ashcraft, 1982, 1987; Campbell, 1987a, 1987b; Campbell & Graham, 1985; Miller, Perlmutter, & Keating, 1984). Arithmetical facts would appear to be stored in the form of associations between digits (in pairs) and answers. When adult subjects solve a problem, the digits are activated, and this activation spreads as far as the associated response (or responses). The speed and precision of the response would then depend on the strength of the association between the digits and answers and the amount of interference (Ashcraft, 1992; Siegler, 1986; Zbrodoff, 1995).

In contrast, children (i.e., from age 5 to 8) appear, for the most part, to use a variety of counting procedures (Baroody, 1987; Carpenter & Moser, 1984; Fuson, 1982). Groen and Parkman (1972) have demonstrated that the reaction time (RT) required for first graders to produce a response to addition problems with a sum less than or equal to 9 increased as a function of the size of the smaller addend (*min* proce-

dure). They suggested that first graders solve simple addition problems by using increments. For example, the mechanism used to solve the problem $3+5$ is a count-up from the larger addend (i.e., 5) through repeated increments of 1, until the number of steps equals 3, or the *min* term (Groen & Resnick, 1977).

The data gathered from studies of young children suggest that development of the ability to add takes the form of a transition from a systematic algorithmic computation to memory retrieval processes. It has been thought that up to the age of 10, in the majority of cases, children count before moving on to the subsequent memory retrieval strategy (Ashcraft & Battaglia, 1978; Ashcraft & Fierman, 1982; Ashcraft & Stazyk, 1981; Svenson & Broquist, 1975). Various empirical arguments have been advanced to support this theory (Fuson, 1982; Geary & Burlingham-Dubree, 1989; Siegler & Shrager, 1984). But things are not that simple.

Solving Additions and Subtractions: From Algorithmic Computation to Memory Retrieval

The performance of young children shows that the increase in RT as a function of the size of the smaller addend is observed only for non-tie additions, whereas almost no size effect is observed in the case of tie additions (e.g., $2+2$; $3+3$) (Ashcraft & Fierman, 1982; Groen & Parkman, 1972; Hamann & Ashcraft, 1985).

Moreover, the observations and analyses undertaken by Siegler (1987; Siegler & Shrager, 1984) have shown that children of kindergarten age, as well as first and second graders, use the *min* procedure for only a third of additions. They retrieve the results of another third of the problems directly from memory and make use of other procedures to solve the final third. Thus children, even at an early age, seem to possess a variety of procedures for solv-

The authors thank Pierre Perruchet, Jane Zbrodoff, and J. R. Anderson for their helpful comments on previous versions of this article. We also wish to thank Sylvie Desportes for assistance in conducting the experiments. Finally, we would like to express our gratitude to Tim Pownall. Correspondence may be sent to either author at the Laboratoire d'Etude des Apprentissages et du Développement, ESA CNRS 5022, Université de Bourgogne, Faculté des Sciences, 6 Boulevard Gabriel, 21000 Dijon, France (e-mail: barouil@satie.u-bourgogne.fr or fayol@satie.u-bourgogne.fr).

—Accepted by previous editor, Geoffrey R. Loftus

ing additions. They select these procedures in an adaptive way—as a function of their probability of success, of the time taken for problem solving, and of the cost of implementing them (Siegler, 1986). These analyses suggest that as a result of practice and feedback, children develop associations between problems (e.g., $4+2$) and potential responses (e.g., $4\ 5\ 6\ 7$). In this way, the distribution of associations may be restricted (i.e., the link between the operation and the correct response would become dominant) and children might then be able to retrieve the correct response rapidly from memory. To summarize, the data show that children retrieve the results from memory at a very early age, and that 1 by 1 incremental counting is the less favored procedure, even among 5- to 6-year-olds. In contrast, research has shown that adults do not systematically retrieve responses from memory for simple addition problems (LeFevre, Sadesky, & Bisanz, 1996; Svenson, 1985). Geary and Wiley (1991) have shown that in 12% of the cases in their study, young adults used verbal counting or a decomposition procedure. Moreover, the use of backup strategies increased with the level of difficulty of the operations.

Overall, the data from the performance of adults and young children in tasks requiring the resolution of simple addition problems show that: (1) at all ages, subjects possess a variety of strategies and, at a minimum, counting and direct retrieval from memory; (2) development does not take the form of the substitution of one strategy (i.e., retrieval) for another (i.e., counting) but instead consists of a change in the *probability* that one strategy will be used instead of another. The youngest subjects primarily use a counting strategy, whereas adults primarily, but not exclusively, retrieve results directly from memory.

The data presented above concern addition. We possess much less information about subtraction (Siegler, 1989; Siegler & Shrager, 1984; Steinberg, 1985; Svenson & Hedenborg, 1979; Wood, Resnick, & Groen, 1975). Despite this, the general observations agree with those made in the case of addition. Subtraction problems are solved initially by manipulating objects and subsequently by verbal counting. Verbal counting strategies for subtraction (in particular, counting down) are more difficult to implement than the verbal counting strategies for addition (e.g., counting up). In effect, subjects are less successful in counting n steps backward than in counting n steps forward (Fuson, Richards, & Briars, 1982). The errors made by subjects using counting strategies are of the same type as those observed in the case of addition (Ilg & Ames, 1951).

The resolution of simple subtraction problems through the direct retrieval of the responses associated with the operands (e.g., $6 - 2 = 4$) is thought to occur at a later stage than that for addition, but chronometric studies provide complex data for interpretation (Siegler, 1989; Svenson & Hedenborg, 1979). Geary, French, and Wiley (1993) have discovered a very high level of variation in the proportions of subjects who retrieve subtractive number facts from memory (between 30% and 80%, depending on the age of the subjects). The detection of subtraction errors

that are characteristic of memory retrieval, such as cross-operation errors (e.g., providing the additive response, 8, to the subtractive problem $6 - 2$), or the existence of a problem-size effect, is not sufficient to allow us to state that the resolution of subtraction problems does not primarily depend on an algorithmic procedure.

Subtraction is more difficult than addition, possibly because it is practised less frequently (Fuson, 1984). This greater level of difficulty becomes apparent as soon as young children start to use a verbal counting procedure, and it is perpetuated by the fact that there are fewer direct memory retrievals than in the case of addition (Fuson & Kwon, 1984). Subtraction problems may therefore continue to be solved most frequently by means of an algorithmic procedure, even among adult subjects. Subtraction would therefore continue to impose a greater cognitive cost than addition, which is more frequently associated with direct memory retrieval (see below).

To summarize, the development of strategies for the resolution of simple addition problems appears to be characterized by a gradual transition from a counting strategy, which is predominant in young children, to a strategy for the direct retrieval of the results from memory, which is the most common strategy among adult subjects. This transition takes place later and is less clear-cut in the case of subtraction.

Our aim in this study was to provide evidence to support this theory by comparing the performance of second graders and university students in two tasks, one involving number arithmetic and the other alphabet arithmetic, using series of addition and subtraction problems (Compton & Logan, 1991; Klapp, Boches, Trabert, & Logan, 1991; Logan & Klapp, 1991; Zbrodoff, 1995). The number arithmetic task consisted of asking subjects to transform, in their heads, series of one to four numbers by adding or subtracting an operand of the value 1, 2, 3, or 4 (e.g., the series 7, 17, 12 together with the operand $+2$ gives 9, 19, 14). In the second task, alphabet arithmetic, the subjects had to perform the same transformations on a series of letters in order to obtain other letters (e.g., the letter *C* together with the operand $+2$ gives *E*; the series *F*, *D*, *T* together with the operand -3 gives *C*, *A*, *Q*). Comparing the performances of the same subjects in these two tasks, one that necessarily requires algorithmic computation (i.e., alphabet arithmetic) because subjects rarely use the alphabet for counting, and one that may mobilize either algorithmic knowledge or the possibly automatized retrieval of numerical facts, should make it possible to determine the extent to which the solving of addition and subtraction problems is automatized at the two grade levels studied here.

Automaticity and the Problem of Resources

Logan and Klapp (1991) used alphabet arithmetic tasks to study automaticity in adults. In effect, alphabet arithmetic is to adults what number arithmetic is to children. It therefore makes it possible to study the strategies used by novices. The authors asked subjects to check the correctness of problems such as $C+2 = E$ (true response) as op-

posed to $C+2 = F$ (false response) (Compton & Logan, 1991; Logan & Klapp, 1991). The subjects had to evaluate their responses as rapidly as possible (their response times [RTs] were recorded) and also state whether they had solved each problem by counting or memory retrieval. The results showed that at the start of the trial, subjects worked by counting through the alphabet: Their verification time was a linear function of the size of the number to be added. They were thus able to mobilize the same types of procedure based on letter-line associations as those of children who use number-line associations (LeFevre, Kulak, & Bisanz, 1991). After extended practice, the subjects claimed to retrieve the results directly from memory, and as a result, the slope of the function linking the RTs to the size of the addend was very close to zero. Klapp et al. (1991) suggested that the transition from counting to retrieval corresponds to the same processes as those that underlie the same transition in children in the case of numbers. Consequently, these experiments simulated children's learning of number arithmetic.

In a second series of experiments, Klapp et al. (1991) used the same tasks in order to study the sensitivity of performance to the (potential) interference caused by competing tasks. They considered that performance becomes automatic when memory retrieval replaces algorithmic counting and that the effect of the size of the addend on evaluation time disappears. The results suggested that subjects can perform arithmetical operations at three levels. At Level I (novice), the verification of the operations takes the form of step-by-step counting through the alphabetical sequence: It is slow, and its duration increases as a linear function of the size of the addends. Moreover, it is easily disrupted by a simple secondary interference task (e.g., the repetition of the same word). At Level II (automatized), the verification times become rapid and independent of the size of the addends. A secondary, repetitive task no longer produces any interference. According to Logan and Klapp (1991), memory retrieval has now replaced the algorithmic procedure. Nevertheless, at this level a secondary task necessitating concurrent retrievals from memory (e.g., listing the months of the year) impairs performance. At Level III (beyond automaticity), verification times are even shorter than in Level II, and no further interference is detected as a result of the secondary task, even when this task is costly.

These data demonstrate that the learning of arithmetical operations takes the form of a transition from algorithmic counting to rapid, direct retrieval from memory, which becomes increasingly fast and insensitive to interference. Logan (1988b) distinguishes between this approach to automaticity (i.e., automaticity-as-memory view) and the approach that juxtaposes automatic processes and cognitively costly processes (Hasher & Zacks, 1979; Posner & Snyder, 1975), and he lists his reservations about the notion of resources. According to him, this concept has particular trouble explaining the development of automaticity and the fact that automaticity does not completely eliminate interference resulting from a secondary

task. However, Logan (1988a) also suggests that the two approaches are not incompatible (Brown & Carr, 1989).

Indeed, a number of recent models combine the idea of cognitive resources with the ideas of activation and memory retrieval (Anderson, Reder, & Lebiere, 1996; Cantor & Engle, 1993; Conway & Engle, 1994; Engle, Cantor, & Carullo, 1992; Just & Carpenter, 1992). The model of working memory proposed by Anderson et al. (1996), which is based on the ACT-R model (Anderson, 1993), seems to be particularly well suited to formalizing problems relating to the limitation of cognitive resources and the automation of retrieval processes.

ACT-R is a model in which a production system acts on declarative memory. Declarative knowledge (e.g., $3+4 = 7$) is viewed in terms of a schematic structure (chunk) that specifies the category to which the knowledge belongs (e.g., additive fact number) and that contains units that encode the contents of this knowledge (i.e., 3, 4, sum = 7). Items of knowledge can be activated by the production rules if they correspond to the goals that form part of these rules (e.g., calculating a sum) and to the data that are to be processed during the current task (i.e., 3 and 4). According to the ACT-R theory, the activation of a chunk depends on the sources of activation which it receives from the elements which are currently present in the field of attention and the strength of the associations between these elements and the chunk. The activation A_i of a chunk is specified by the equation $A_i = \sum W_j S_{ij}$, where W_j represents the level of activation of the elements for processing and S_{ij} is the strength of the association between these elements and the chunk. The probability that an item of knowledge will be retrieved from long-term memory is an exponential function of its activation A_i ; the access time is a negative exponential function of this activation. As the level of the activation received rises, the probability that the knowledge will be activated increases, and the access time falls.

Anderson et al. (1996) insist on the difference between activations, which control retrieval from declarative memory, and the sources of activation, which reflect the salience of an element in the situation in which attention is focused on it. The authors believe that what is limited is not so much the maximum level of activation, which depends in part on the strength of the associations S_{ij} , but the total quantity of the sources of activation (i.e., the attentional resources: $\sum W_j$). Thus, when subjects simultaneously have to perform two tasks that both require the allocation of attentional resources, the total quantity of the sources of activation is shared between the elements that have to be processed in each of the two tasks. The subsequent limitation of each of the sources of activation would therefore have a direct impact on the activation of the items of knowledge and, consequently, on their accessibility.

In order to demonstrate that the limitation of working memory capacity is due to memory retrieval problems, Anderson et al. (1996) followed Carlson, Sullivan, and Schneider (1989) in using a dual-task paradigm in which subjects had to retain a series of numbers (of varying span) in memory while solving an equation (of varying diffi-

culty). After solving the equation, the subjects were asked to recall the series of numbers. The authors observed that an increase in the memory span resulted in longer problem-solving times and that this effect increased as the problems became more difficult. The distribution of sources of activation to a growing number of elements that form part of either the memory load or the problem that is to be processed makes it possible to explain the memory span effect in terms of the mechanical reduction of the activation levels.

The ACT-R model may allow us to conclude that automaticity, as described by Logan (i.e., through direct memory retrieval) is cognitively less costly than implementing a counting algorithm. In effect, an algorithmic procedure for adding n requires n successive memory retrievals (in the alphabet or number chain) plus the monitoring of the number of steps performed, thus resulting in a greater cognitive cost. This model also makes it possible to explain why at Level 2 (automatized) as defined by Klapp et al. (1991), sequential month saying interferes with the verification of alphabet arithmetic. The memory retrieval of the list of words would require the allocation of resources, and this would diminish the level of activation of the result that is to be verified. Similarly, training beyond what is needed for automaticity (i.e., direct retrieval) would lead to the strengthening of the link between the operands and the result (e.g., A , 2, and C when the problem for verification is $A + 2 = C$). Activation of the result (i.e., C) beyond the threshold required for retrieval would then require the allocation of a lower level of resources, and the interference caused by sequential month saying would disappear.

To summarize, Anderson's ACT-R model allows us to integrate Logan's automaticity-as-memory view with the conception of automaticity as a reduction of cognitive load. In our experiment, subjects had to resolve series of operations on either numbers or letters, while maintaining the results in memory. This paradigm is akin to the task used by Anderson et al. (1996), and the same phenomenon (i.e., the effect of memory load and problem difficulty) should result from the tradeoff between processing and storage.

The Present Study

The aim of our research was to confirm that algorithmic procedures dominate in the solving of addition and subtraction problems (with a sum varying between 2 and 26) in children (second graders), whereas adults (university students) most frequently solve the same problems by means of the direct retrieval of the results from memory.¹ Thus, when subjects have to apply an operation (e.g., +3) to a series of numbers (e.g., 13, 7, 18, 22) while retaining the result of each transformation in memory, it should become more difficult to retain the results as the cognitive cost of the successive calculations increases.

We assume that the algorithmic resolution of a problem is cognitively costly and that this procedure limits the number of items of information that can be stored or processed in working memory. In contrast, the cost of resolution through direct memory retrieval is much lower and there-

fore places fewer constraints on the number of items that can be stored and/or processed in working memory. In consequence, the effect on performance of the number of items to be transformed and results to be retained in memory (item span, IS) should decrease as processing becomes increasingly automatized. Moreover, the effect of the size of the operand (operation span, OS) should be more marked in the case of algorithmic strategies than for direct memory retrieval. In effect, OS determines the number of steps that subjects have to perform and monitor during computation when they employ an algorithmic strategy (e.g., $8 + 4$: 9, 10, 11, 12; result, 12). In contrast, memory retrieval should not be sensitive to OS (Klapp et al., 1991; Logan & Klapp, 1991).

The hypothesis that there is a transition from algorithmic processing to direct memory retrieval predicts that children should employ an algorithmic procedure when solving addition and subtraction problems involving the numbers 1–26 as well as the 26 letters of the alphabet. In consequence, performance patterns should be approximately equivalent (or at least very similar) for letters and numbers, additions and subtractions. This is because the systematic use of counting algorithms and the rare intervention of direct retrieval should mean that there is no, or only a low level of, interaction between the type of material for transformation (numbers vs. letters) and the OS or IS on the one hand and, on the other, between the operations (addition vs. subtraction) and the OS or IS.

In contrast, adults should only systematically apply an algorithmic procedure in the case of alphabetical problems (given that the subjects are tested during a single session only and are therefore unable to achieve a sufficient level of automaticity, unlike in Klapp et al., 1991). In consequence, performance in number arithmetic, which is probably heavily based on direct retrieval, should differ considerably from the performance observed for alphabet arithmetic. These differences should also be more accentuated for addition than for subtraction. We predicted interactions between the IS and the material, on the one hand, and the OS on the other. In effect, since the adults rely very heavily on algorithmic procedures for alphabet arithmetic and direct retrieval for number arithmetic, their letter performance should be significantly weaker than their number performance. Furthermore, this decline in performance as OS and IS increase should be more accentuated in the case of letters (i.e., when the use of the algorithm imposes a greater cognitive cost) than in the case of numbers. Finally, performance differences between numbers and letters should be greater for addition problems (greater automation with numbers) than for subtraction problems (more frequently solved using an algorithm).

Anderson's model predicts an interaction between OS (i.e., the difficulty of the problem) and IS. This interaction should be more marked in the 2nd-grade subjects than in the 12th-grade subjects. Two factors may contribute to this phenomenon. The first is the possible growth in attentional resources with age, which may result in a reduction

of the IS and OS effects. The second is the automation of the task of verbally working through the alphabet or performing numerical calculations. As far as the alphabet is concerned, automation would result from the strengthening of associations between successive terms. Thus the term $n+1$ would reach the threshold necessary for its retrieval more rapidly and easily the stronger its association with the term n . Consequently, the cognitive cost of transformations should be less for the adults than for the children, and the former should be less sensitive than the latter to variations in OS.

In short, adult performance on letter problems should resemble children's performance on number problems.

METHOD

Subjects

Twenty children in 2nd grade (mean age = 7,5 years, min = 6,11, max = 7,10) and 20 adolescents in 12th grade (mean age = 17,5 years, min = 16,4, max = 19,4) took part in the experience. There were 10 females and 10 males in each group.

Material

The material took the form of cards measuring 15×10 cm on which were printed series of letters or numbers 12 mm tall. These series contained from one to four items (e.g., a series of four letters could be CKQH; a series of four numbers, 13-18-5-9). The number of items constituted the item span. To the right of the series were an operator sign (+ or -) and an operand of the value 1, 2, 3, or 4. The latter two specifications indicated the direction and size (span) of the operation that the subjects had to apply to each of the items in the series. In the case of the letters, the signs + and - referred to the direction in which subjects had to move through the alphabet (i.e., forward for + and backward for -). The operand (1-4) indicated the number of steps to be performed. For example, the series *AHT* transformed through the operation +2 gives *CJV*. In the case of the numbers, the signs referred to addition and subtraction operations. Thus the series 3-18-9 transformed through the operation +2 gives 5-20-11. The letter series were formed from the 26 letters of the alphabet and the number series from the whole numbers 1-26.

Four letter series, one for each IS (from 1 to 4), were formed for each OS (from 1 to 4) and each direction (+ or -). The series were formed at random, in accordance with the following constraints: (1) The same letter was not allowed to occur in two series assigned to the same operation; (2) letters requiring impossible transformations were removed (e.g., $A-1$ or $Y+3$); (3) the result of the operation was not allowed to contain more than one letter from the original series; (4) any series containing two successive letters was removed. For the letters, a total of 8 sets (4 operation spans \times 2 directions) of 4 series (span of 1-4) were formed. The number series were then formed by substituting the number corresponding to the position in the alphabet of each letter in the letter series (e.g., 6-24-20 for *FXT*). The material therefore contained a total of 64 series (4 item spans \times 4 operation spans \times 2 directions \times 2 types of material: letters or numbers). These 64 series were divided into four blocks, which were determined by combining the direction of transformation with the type of material.

Four presentation sequences were constructed for these blocks (Latin square). In each of the two age groups, 5 subjects were assigned to each of the four sequences. Within each block, the series were presented in ascending order of OS (i.e., 1, 2, 3, and 4) and, for each OS, by ascending order of item span. For example, for the letter + block, the subjects were presented with the series of one, two, three, and then four letters assigned to +1 and then the four-letter series assigned to +2 and so on.

Procedure

The trial took the form of an individual session conducted in a quiet school room. The subjects were instructed that the task consisted of transforming numbers or letters by addition or subtraction. In the case of the letters, the experimenter explained that the operations could be performed by considering the alphabetical order as a numerical order and gave an example for each direction of transformation. Two sets of four series consisting of the days of the week or the months of the year were used in a training phase, and the subjects were given feedback until they reached a completely correct answer. All the subjects perfectly understood the principle of transforming the nonnumerical series.

The experimenter then presented the cards specifying the series for transformation and the associated operations one at a time. The subjects performed these operations in their heads while continuing to view the series for transformation. When they thought that they had finished, they told the experimenter, who then turned the card face down on the table as the subject wrote his or her response in a notebook. The subjects had to write the result of transforming each item in the order in which these items appeared in the series. Any inversion was considered to be an error (e.g., for *FXT*+2, the response *HVZ* was deemed to be an error, the correct response being *HZV*). If the subjects were unable to transform an item in the series or had forgotten the result, they replaced it with a dash (e.g., $H-V$).

A stop criterion was used in order to shorten the testing session. For a given OS, the series were presented in increasing order of size (one, two, three, four items). When, for a given IS, subjects replaced one or more items with a dash, the series with larger item spans were not presented and were scored as failures. In contrast, presentation was not interrupted if the subjects made an error. For example, following the response $H-V$ to the item *FXT*, the four-letter series was not presented. In contrast, presentation was not suspended following the response *HUV* to the same series, even though this response was incorrect. We proceeded in this way because we considered that the absence of a response indicated that the subject's processing capacities were saturated at this size of transformation, whereas an incorrect response could be due to a simple computation error. Indeed, a preexperiment had shown that when subjects failed to transform a series for a given OS (forgetting a result), they also failed systematically on larger series. The dependent variable was the number of series transformed correctly.

RESULTS

Each subject solved a maximum of 64 problems. The global success level for all ages combined was 65.4%. Table 1 gives the number and percentage of errors observed (series not transformed correctly) in each experimental condition. As in the table, the percentages specified in the remainder of the text are percentages of errors.

A 2 (grade: 2nd and 12th) \times 4 (IS: from 1 to 4) \times 4 (OS: from 1 to 4) \times 2 (operation: addition and subtraction) \times 2 (material: letters and numbers) analysis of variance was performed with repeated measures on the last four factors (see Tables A1 and A2). All the main effects were significant. The 2nd-grade subjects made more errors (49.7%) than did the 12th graders (19.6%) [$F(1,38) = 77.54, MS_e = .751, p < .001$]. The series of numbers were easier to process (24.3%) than the series of letters (45.1%) [$F(1,38) = 119.80, MS_e = .231, p < .001$]. The proportion of errors increased both with IS (8%, 24%, 45%, and 63% of errors for the spans 1, 2, 3, and 4, respectively) [$F(3,114) = 186.44, MS_e = .198, p < .001$], and with the OS (21%, 33%, 42%, and 43% for the spans 1, 2, 3, and 4)

Table 1
Numbers of Errors Observed for Each Problem,
With Percentages of Errors, as a Function of Grade

Grade	Operation Span	Item Span										
		Numbers					Letters					
		1	2	3	4	% Error	1	2	3	4	% Error	% Error
Addition												
1	1	0	1	3	16	25	0	1	6	17	30	28
	2	2	1	11	18	40	0	9	16	19	55	48
	3	0	5	11	18	43	2	10	18	20	63	53
	4	0	5	15	19	49	2	11	17	20	63	56
	% Error	2	15	50	89		5	39	71	95		
Subtraction												
1	1	0	1	10	14	31	0	5	7	19	39	35
	2	0	5	13	17	44	0	7	16	18	51	48
	3	2	9	14	20	56	7	14	19	19	74	65
	4	3	10	15	19	59	8	16	18	19	76	68
	% Error	6	31	65	88		19	53	75	94		
Total % Error	4	23	58	88		12	46	73	94			
Addition												
12	1	0	0	0	2	3	1	3	4	5	16	9
	2	0	0	0	1	1	2	0	6	9	21	11
	3	0	1	1	5	9	4	7	9	11	39	24
	4	0	0	1	2	4	3	5	6	10	30	17
	% Error	0	1	3	13		13	19	31	44		
Subtraction												
12	1	1	1	1	0	4	2	2	8	7	24	14
	2	0	0	1	5	8	0	5	13	16	43	25
	3	0	1	0	4	6	4	5	12	14	44	25
	4	0	2	2	3	9	6	9	15	14	55	32
	% Error	1	5	5	15		12	26	60	64		
Total % Error	1	3	4	14		14	23	46	54			

$[F(3,114) = 54.87, MS_e = .115, p < .001]$. Finally, additions resulted in fewer errors (30.5%) than did subtractions (38.8%) $[F(1,38) = 22.87, MS_e = .192, p < .001]$. These effects were significant within each grade considered separately.

As predicted by the hypotheses, the level \times material and level \times material \times IS interactions were significant. The effect of the type of material for processing was weaker among 2nd-grade subjects (56% errors on letters, 43% errors on numbers) than among 12th-grade subjects (34% errors on letters, only 5% errors on numbers) $[F(1,38) = 16.93, MS_e = .231, p < .001]$. At both levels, the number of errors increased as a function of IS. However, the interaction between the material and the IS took different forms, depending on the level of the subjects $[F(3,114) = 7.81, MS_e = .153, p < .001]$. In the case of the 2nd-grade subjects, the difference between numbers and letters did not increase regularly with the increase in IS (differences of 8%, 23%, 15%, and 6% for spans of 1, 2, 3, and 4). In contrast, an increase in this difference was observed among the 12th-grade subjects (13%, 20%, 42%, and 42%).

These effects suggest that when OS was held constant, the increase in IS (i.e., the working memory load due to the increase in the information to be stored) affected young children's calculations of number and letters to a very similar extent. The material \times IS interaction was due

to an increase in the number/letter difference between spans of 1 and 2 as well as to a reduction in this difference for spans of 3 and 4. The form of this interaction did not allow us to reject the hypothesis that the two types of material are processed in the same way. Indeed, the IS effect was extremely strong for both letters and numbers. In contrast, 12th-grade subjects were largely insensitive to the increase in IS in the case of numbers (1%, 3%, 4%, and 14% errors for ISs of 1, 2, 3, and 4, respectively), whereas increasing this span resulted in a marked decline in performance in the case of letters (14%, 23%, 46%, and 54% errors for ISs of 1, 2, 3, and 4, respectively). The form of this interaction suggests that the adults were using a more costly strategy for the letters than for the numbers.

The OS effect was large for both types of material at grade 2 and, as predicted, the material \times OS interaction was not significant $[F(3,57) = 1.98]$. In contrast, the OS effect was observed for the letters only at 12th grade. At this level, the OS effect on numbers was practically zero (3%, 4%, 7.5%, and 6% errors for spans 1, 2, 3, and 4, respectively; the linear tendency was not significant, $F < 1$), whereas there was a strong effect on letters (20%, 32%, 41%, and 43% errors for spans 1, 2, 3, and 4, respectively) and the material \times OS interaction was significant $[F(3,57) = 6.47, MS_e = .092, p < .001]$. The weak OS and IS effects on the resolution of series of numbers suggests that the 12th-

Table 2
R² Values (Percentage of Experimental Variance
Explained by Each Source) Obtained as Part of the
ANOVAs Presented in Table A2 for the Various
Sources of Variation as a Function of Grade

Sources	Grade 2	Grade 12
Material	.003	.431
Operation	.012	.040
Operation span (OS)	.110	.059
Item span (IS)	.759	.225
Material × Operation	.000	.020
Material × OS	.004	.030
Material × IS	.008	.083
Operation × OS	.005	.018
Operation × IS	.007	.015
OS × IS	.034	.030
Material × Operation × OS	.001	.004
Material × Operation × IS	.003	.014
Material × OS × IS	.011	.009
Operation × OS × IS	.007	.016
Material × Operation × OS × IS	.007	.009

grade subjects were using a strategy that differed from that which they employed for letters as well as from that used by the 2nd-grade subjects for numbers and letters. This strategy, which is insensitive to OS, could be the direct retrieval of the result from memory (i.e., $16+3=19$). Accordingly, the absence of an OS effect is interpreted by Logan and Klapp (1991) and Klapp et al. (1991) to be an index of the automatic direct retrieval of results from memory.

The operation type effect was significant and, consistent with our hypothesis, the operation did not interact with the material to be transformed in grade 2 ($F < 1$), whereas this interaction was significant in adults [$F(1,19) = 4.87$, $MS_e = .244$, $p < .05$]. In grade 2, subtractions elicited more errors than additions, but this difference was the same for letters (60% and 53% errors for subtraction and addition, respectively) and numbers (48% and 39%). In grade 12, the effect of operation was null for numbers (7% and 4%), whereas subtraction elicited more errors (41%) than did addition (27%) on letters. The disappearance of the operation effect on numbers suggested that adult subjects used the same strategy for both the additions and subtractions—namely, a direct retrieval from memory. As a result, both the IS and the OS effects were weak for subtraction on numbers (see Table 1).

As far as letters are concerned, the operation effect interacted weakly with IS for 2nd-grade subjects (5%, 39%, 71%, and 95% errors for addition, and 19%, 53%, 75%, and 94% for subtraction for IS of 1, 2, 3, and 4, respectively). The cognitive cost associated with the processing employed by subjects was greater in the case of subtraction than addition; but this difference remained approximately constant whatever the IS, and the interaction was probably the result of a ceiling effect. In contrast, the IS effect was stronger for subtraction (12%, 26%, 60%, and 64% errors for ISs of 1, 2, 3, and 4, respectively) than for addition (13%, 19%, 31%, and 44% errors for ISs of 1, 2, 3, and 4, respectively) in grade 12, and the operation × IS × level interaction was significant [$F(3,114) = 4.91$, $MS_e = .141$,

$p < .01$]. The results suggest that the 12th graders continue to employ a step by step strategy for working through the alphabetical sequence. However, the differential effect of the operation as a function of IS may result from the fact that the age-related automation of the alphabetical sequence is more pronounced in the forward than in the backward direction.

The hypothesis of a tradeoff between processing and storage predicts two interactions: IS × OS and level × IS × OS. As predicted, the IS × OS interaction was significant at both levels separately [for grade 2, $F(9,171) = 5.44$, $MS_e = .106$, $p < .001$; for grade 12, $F(9,171) = 2.02$, $MS_e = .099$, $p < .05$] or globally [$F(9,342) = 4.21$, $MS_e = .106$, $p < .001$], and the level × IS × OS interaction was significant [$F(9,342) = 3.49$, $MS_e = .106$, $p < .001$]. The OS effect increased with increases in IS. This was particularly clear for 2nd-grade subjects with an IS of 1 (from 0% errors for OS1 to 16% errors for OS4), an IS of 2 (from 10% for OS1 to 53% for OS4), and an IS of 3 (from 33% for OS1 to 81% for OS4). Subjects almost never succeeded in the four-item tests whatever the OS (83%, 90%, 96%, and 96% errors for OS1, 2, 3, and 4, respectively), whether they were dealing with numbers or letters. According to the hypothesis that this tradeoff occurs in a resource pool the capacity of which grows with age, the slopes due to the increase in OS were shallower for the 12th-grade subjects. The OS effect (i.e., the increase in percentage errors from OS1 to OS4) was 6% on IS1, 12% on IS2, 14% on IS3, and 16% on IS4.

The results suggest that adults employ different processing strategies for numbers and letters, whereas identical strategies are used by 2nd-grade subjects. This results in the variability of the respective importance of each of the sources of variation as a function of age (Table 2).

The factor that has the greatest effect on the performance of 2nd-grade subjects is the IS ($R^2 = .759$). This is followed by the OS ($R^2 = .110$). The factor of material is of minimal importance ($R^2 = .033$) as are the IS × material (.008) and OS × material (.004) interactions, even though the effects of these three factors are significant. The effect of the type of material is weak in comparison with the IS and OS effects. In contrast, the performance of adult subjects is primarily affected by the type of material for processing (.431), while the importance of IS falls considerably at this level (.225). The change from numbers to letters has only a small impact on the performance of 2nd-grade subjects but a considerable impact on 12th-grade subjects (for whom the IS × material interaction is the third largest source of variation [$R^2 = .083$], with all the simple interactions involving the factor of material having a considerably greater effect than at grade 2).

To summarize, the results confirmed the main hypotheses underlying this experiment. Processing and storage activities are in competition for a single resource pool, thus leading to an OS × IS interaction in particular when processing imposes a high cognitive load (i.e., at 2nd grade). At this level, there appears to be no difference between the strategies used for processing the alphabetical and nu-

Table 3
R² Values From the Application of Anderson's
ACT-R Model and Multiple Regression Slopes Between
Operation Span (OS) and Item Span (IS) on the One Hand
and, on the Other, the Probability of Error Observed
as a Function of the Material and Grade

Grade	Material	R ²	Slope	
			IS	IS × OS
2	Numbers	.879	.206	.031
	Letters	.922	.184	.037
	Overall	.916	.197	.034
12	Numbers	.607	.025	.006
	Letters	.864	.076	.027
	Overall	.863	.050	.017

merical sequences. It is therefore likely that subjects perform transformations by working step by step through both the numerical and alphabetical sequences and monitoring the number of steps that they have completed. This results in a major OS effect.

Although the letter processing strategy appears to be the same at the 12th-grade level, the same is not the case for numbers. The error level on number problems no longer increases as a function of OS, and processing is sufficiently automatized for IS to have only a weak effect on performance. Since only minimal resources are allocated to processing, retention in memory and the retrieval of the results pose few problems. The error level was very low, peaking at 15% (subtraction, IS 4) as against 88% in the same condition at grade 2. Finally, we observed an automation of the alphabetical chain during the course of development, and the IS and OS effects were consequently weaker in 12th-grade than in 2nd-grade subjects.

The facts observed here help us explain three problems. The first is related to the evaluation of the model of a tradeoff between processing and storage. The second is related to the reasons behind the superiority of 12th-grade over 2nd-grade subjects. It is possible to point to two reasons for this superiority. The first is the automation of processing. Since the cognitive cost of processing is less, there is more space available for storage, and this leads to improved performance. This hypothesis does not require us to assume that the resource pool also grows with age (Case, 1985). The second reason may be that, alongside automation and the associated effects, there is growth in the pool of cognitive resources (Halford, 1993; Pascual Leone, 1970, 1988). The superiority of 12th-grade subjects would then be due to the cumulated effects of an increased volume of resources and greater automation. The third problem concerns the relationship between automaticity and cognitive resources. We shall analyze these three problems in the light of the ACT-R model.

DISCUSSION

Anderson et al. (1996) suggest that their model may be able to account for the limitation of a single resource pool within a processing unit such as the central processor pro-

posed by Baddeley (1986). We can use this model directly to investigate the tasks studied here. The subjects in our experiment were asked to transform series of items (letters or numbers) and store the results of these transformations. Here the transformations play the same role as equation solving in the Anderson et al. (1996) experiment, while the size of the transformation (i.e., from 1 to 4) corresponds to the difficulty of the problems. The results that the subjects have to retain in memory correspond to the memory load, and the item span is the equivalent of the lists for memorization.

Evaluation of the Tradeoff Model

The ACT-R model stipulates that the probability of a correct response is a function of both the number of operations that have to be performed (i.e., the number of steps by which each element has to be modified, which is dependent on OS) and the number of items that have to be retrieved from memory. Anderson et al. (1996) suggest that the probability $P(s)$ of finding the solution is given by the equation

$$P(s) = (P_R)^n \times (P_p)^m,$$

where P_R is the probability of retrieving an item, n is the number of items, P_p is the probability of correctly implementing a procedure (here, one step in the sequence), and m is the number of procedures to be implemented. The probability of success is an exponential function of the number of items to be retrieved and the number of procedures to be implemented, and the log probability of success is therefore a linear function of these parameters. In consequence, the model can be simply tested by using a multiple linear regression on log probabilities of success.

Concerning the variable n , we shall restrict ourselves to the number of results that have already been calculated and that have to be retrieved from memory in order to produce the final response (i.e., IS). Of course, processing itself requires retrievals from long-term memory both in the case of a step by step (i.e., given item x in the sequence, what is item $x + 1$?) and a direct memory retrieval strategy (e.g., $9 + 4 = ?$). However, these retrievals and the difficulties that they may cause are already included in the OS in cases where subjects use a step by step algorithmic strategy.

The variable m is the number of times that the step by step procedure has to be used in a given test. We assume that each step in a series requires the firing of the procedure. The value m , which is dependent on IS and OS, is therefore IS × OS (e.g., if IS = 2 and OS = 3, subjects have to use the procedure for moving one step in the relevant series six times). The slopes associated with each of these variables will provide an estimation of the probability of the failure to retrieve a result that is stored in memory and an estimation of the probability of the failure to implement a procedure for moving within the (alphabetical or numerical) sequence that is to be processed. These slopes are therefore indicators of the cognitive cost incurred by the retrieval of an item and the movement of one step through the sequence. Table 3 gives (1) the values of R^2 for

the multiple linear regression testing the ACT-R model, where $n = IS$ and $m = IS \times OS$; and (2) the slopes associated with the IS (i.e., the probability of the failure to retrieve a result) and $IS \times OS$ parameters (i.e., the probability of the failure to implement a procedure for moving one step in the relevant series) as a function of the material for processing and the level.

Anderson's model has proved to be particularly powerful by accounting for 92% of variance on the letter problems and 88% of the variance on the number problems at grade 2. It should be noted that at this level, the IS and $IS \times OS$ slopes for the numbers and letters are very similar. The addition of an item to the series for transformation resulted in an increase of approximately .20 (.209 and .184) in the probability of error, and the system was saturated at ISs of 4. The retention and retrieval of 4 memory items left only a very low level of resources available for computation, even in the case of OSs of 1, thus resulting in the observed ceiling effect.

In contrast, the slopes for the 12th-grade subjects are much shallower and the model accounts only poorly for the performances observed for numbers ($R^2 = .607$), although it continues to explain letter performance ($R^2 = .864$). As seemed likely from a simple observation of the results, this suggests that the strategy that these subjects use when transforming series of numbers is different from that used for letters.

We have suggested that OS no longer has any effect on number computations at this level because subjects retrieve the results directly from memory and the associated cognitive cost is no longer a direct function of the size of the OS. The fall in the $IS \times OS$ slope for letters between the two levels (from .037 to .027) may point to an automation of the alphabetical sequence, primarily in the canonical (forward) direction.

Superiority of Adults to Children

The regression analysis also made it possible to give an answer to the second question—namely, that relating to the reasons for the better performances observed among the 12th-grade subjects. Can this be explained simply in terms of the automation process, which frees up resources that are then allocated to storage requirements, or is there a simultaneous increase in total processing capacity? We have suggested that the slopes associated with $IS \times OS$ represent an estimation of the parameter P_p in Anderson's model (or, more precisely, $1 - P_p$, since we are measuring errors rather than successes). This probability that the step by step procedure will fail is an indicator of the cognitive cost of moving one step in the numerical or alphabetical sequence.

The data indicate that at grade 2 and for the numbers, the slope associated with OS (.031) was similar to that observed for letters at grade 12 (.027). If our analysis is correct, it follows that the cognitive cost of moving through the number sequence at 2nd grade is comparable to that of moving through the alphabetical sequence at 12th grade. In other words, children in the 2nd grade appear to use the

number sequence as a simple, ordered series that they can work through in either direction by monitoring the number of steps that they have completed in the same way as 12th-grade subjects perform calculations within the alphabetical sequence. Thus, the cognitive load imposed on 12th-grade subjects asked to perform calculations within the alphabetical sequence is similar to that on younger children who are asked to perform numerical calculations.

This situation is comparable to the one designed by Case, Kurland, and Goldberg (1982), who asked adults to perform a counting span task in which the subjects had to count sets by using a series of arbitrary terms that they had learned earlier. The authors observed that the working memory spans measured for the adults were comparable to those exhibited by 6-year-old children who performed the task while counting the objects by using the normal number sequence. Case (1985) used these results to argue that total processing space remained constant throughout development. When adults' counting efficiency is rendered comparable to that of young children, the space that remains available for the storage of the results should be comparable and the observed span identical.

This argument can only carry conviction if we make certain that the cognitive cost linked to the counting activity is identical in the two groups (i.e., for the children in the case of the traditional sequence of numbers, and for the adults, using random terms learned beforehand). This level of control is possible with the paradigm that we adopted. The $IS \times OS$ slopes for numbers at 2nd grade and letters at 12th grade are almost identical. This suggests that the cognitive cost associated with number arithmetic at 2nd grade is identical to that associated with alphabet arithmetic at 12th grade. If Case's (1985) reasoning is correct, we should expect to observe comparable performances. However, the 12th-grade subjects made far fewer errors on the letter problems (33.9%) than did the 2nd-grade subjects on the number problems (43.3%). Table 3 shows that the $IS \times OS$ slopes are comparable, but that the slope associated with IS is steeper for 2nd-grade subjects in the number condition (.206) than for 12th-grade subjects in the letter condition (.076). Thus, despite the fact that the cognitive cost of processing was comparable, the effect of the number of items to be processed was greater in the younger subjects.

Two hypotheses can be advanced to explain this phenomenon. The first supposes that 12th-grade subjects possess a larger pool of resources, which allows them to store and retrieve more items even when the cognitive cost of processing is identical. According to this hypothesis, our results contradict Case's (1985) hypothesis of constant total processing space during development. Two processes would then account for the improved performance of the 12th-grade subjects: the automation of processing (visible in the letter condition, where the slope associated with OS fell from .037 at 2nd grade to .027 at 12th grade) together with an increase in processing capacity, which would result in improved performance even if processing efficiency remained constant.

In contrast, the second hypothesis is compatible with Case's (1985) model. This holds that in the same way that processing (i.e., calculating results) becomes increasingly efficient and therefore less costly with age, the processes involved in the memory storage and retrieval of information are also subject to improvement or automatization.² In effect, it is possible that the activities of storage and retrieval involve strategies for the retention and retrieval of information. The implementation of these strategies could become automatized with age, or more effective strategies could be used. In both cases, this would result in improved recall performance even if the processing required to obtain the results for memorization involved the same cognitive cost, which, as we have seen, is indeed the case. Thus it might be supposed that adults perform better than children simply because they have automatized the activities necessary for the completion of the task. This automation would relate to both the processing necessary for the calculation of the results and the activities necessary for their retention and retrieval.

In the ACT-R model, the automation of processing can be thought of as a strengthening of the links between the successive terms in the sequence through which the subject is working (numerical or alphabetical). If we assume equal attentional resources, a given stimulus (e.g., the letter *G*) would activate the following stimulus (i.e., *H*) more strongly, and this would lead to improved performance in the calculation of the results. In effect, this greater ease in working through the sequence could free up resources and thus improve the monitoring of the number of steps to be performed. When adults are confronted with number problems, this automation will cause the automatic activation of the result by the term that is to be transformed and the operator that is to be applied to it. This view is compatible with Case's (1985) model. Similarly, a self-repetition strategy would result in a modification in the strength of the links between the various terms for recall. The same quantity of attentional resources would therefore lead to a higher level of activation of the set of terms for recall and thus result in improved performance.

In fact, in Anderson's (1993) model, the level of activation of a given item and the probability that it will be accessed depend both on the quantity of attentional resources and on the strength of the link between this item and the input items. Thus the model is compatible with the hypothesis of improved performance due either to a global increase in capacity (i.e., an increase in attentional resources) or to the strengthening of the associative links in the network, or, of course, to the two factors combined. Nothing in the obtained results allows us to decide between these two hypotheses. Although the second is compatible with Case's (1985) model, it presupposes not only that the efficiency of processing improves but also that the efficiency of all the activities used in solving the task improves.

Automation, Memory Retrieval, and Resources

In our introduction we emphasized that Logan (1988a) suggested that his automation as memory model might be

compatible with models that mobilize the concept of resources. A comparison of the performance in alphabet and number arithmetic at the two ages in question makes this type of synthesis conceivable. Our results clearly indicate that letter and number calculations at 2nd grade and letter calculations alone at 12th grade are accompanied by a high cognitive cost, but that this cost is considerably lower in the case of numbers at 12th grade. This is entirely compatible with the hypothesis of the direct retrieval of results from memory in the case of numbers at 12th grade.

However, Anderson's memory model supposes that memory retrieval, however automatized it may be, requires the allocation of resources. Klapp et al. (1991) observed that at Level III, subjects' RTs were no longer affected by the secondary task of sequential month saying. We observed that IS had an effect on performance at 12th grade, even in the case of numbers. Thus, a competing memory load would appear to affect an automatized process (Fayol, Largy, & Lemaire, 1994; Largy, Fayol, & Lemaire, 1996; Lemaire, Abdi, & Fayol, 1996). However, Table 1 shows that the effect of IS on numbers is primarily due to the IS of 4. This suggests that the automatized process of memory retrieval is not affected unless the competing memory load is high, an observation that can also be derived from Anderson's model. Since the associations are probably very strong (S_{ij}), even weak sources of activation are likely to cause activation of the result at a level that is sufficient for its retrieval. Only a high competing memory load (e.g., IS = 4) is likely to disturb this process. In contrast, the computation strategy used by the 2nd graders for both letters and numbers and by the 12th graders for letters only is sensitive to even the slightest increase in the competing memory load, as is shown by the regular increase in the failure rates as a function of IS in the three situations (see Table 1). This suggests a high cognitive cost.

Moreover, the perception of automation as memory retrieval (Logan, 1988b), together with the hypothesis that learning strengthens the associative links between items in long-term memory (Anderson, 1993), makes it possible to account for the differences that exist between children and adults in the processing of both numbers and letters. In the first case, as predicted by Logan's model, there would be a shift from algorithmic computation to memory retrieval, and this would result in the disappearance of the OS effect and a considerable reduction in the IS effect. Note that these two phenomena affected addition and subtraction in the same way, suggesting that the shift from algorithmic computation to direct retrieval occurs in both operations. In the second case (i.e., letters), the supposition of a similar change of strategy is rendered impossible by the persistence of strong IS and OS effects. However, the automation of the verbal sequence in adults (as a result of the stronger association between successive items) helps to explain why the OS effect is weaker. In this case, it is easier for subjects to work through the alphabet. The reduced cognitive cost that results from this automation would then explain why the IS effect is weaker for letters in the adult subjects.

Thus, by introducing the concept of resources into a memory retrieval model, it is possible to think in terms of two types of automation which may occur during development or learning. One takes the form of the substitution of memory retrieval strategies for algorithmic strategies, as is suggested by Logan. The other appears to affect the algorithmic strategies themselves by permitting the faster and less costly memory retrieval of the items of information that are necessary for the implementation of these strategies as Anderson's ACT-R model suggests. The ACT-R model supposes that the learning mechanism is based on a variation in the strength of the connections between items. When we consider the automation of the alphabetical sequence, it would appear that this model is more appropriate to the instance theory of automaticity put forward by Logan (1988b, 1992).

In conclusion, taken overall, these results confirm the hypothesis that development in solving simple addition and subtraction problems takes the form of a change of strategy, with child subjects moving from an algorithmic strategy of counting up or counting down toward a strategy for the direct retrieval of results from memory at the adult level. The similarity of the performance patterns exhibited by the 2nd graders on numbers and the 12th graders on letters points to the existence of a step by step movement strategy through the sequence of numbers in young children. The considerable effect of the type of material to be processed on adult subjects suggests that this strategy is replaced by a strategy for the direct retrieval of results in the case of numbers but not of letters. As suggested by Logan (1988a), the introduction of the idea of resources into a memory retrieval model provides a promising starting point for the explanation of these phenomena. Nevertheless, the detailed functional analysis of this transition will require further study.

REFERENCES

- ANDERSON, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- ANDERSON, J. R., REDER, L. M., & LEBIERE, C. (1996). Working memory: Activation limitations on retrieval. *Cognitive Psychology*, **30**, 221-256.
- ASHCRAFT, M. H. (1982). The development of cognitive arithmetic: A chronometric approach. *Developmental Review*, **2**, 213-236.
- ASHCRAFT, M. H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In J. Bisanz, C. J. Brainerd, & R. Kail (Eds.), *Formal methods in developmental psychology: Progress in cognitive development research* (pp. 302-336). New York: Springer-Verlag.
- ASHCRAFT, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, **44**, 75-106.
- ASHCRAFT, M. H., & BATTAGLIA, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 527-538.
- ASHCRAFT, M. H., & FIERMAN, B. A. (1982). Mental addition in third, fourth, and fifth graders. *Journal of Experimental Child Psychology*, **33**, 216-234.
- ASHCRAFT, M. H., & STAZYK, E. H. (1981). Mental addition: A test of three verification models. *Memory & Cognition*, **9**, 185-196.
- BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- BAROODY, A. J. (1987). The development of counting strategies for single-digit addition. *Journal for Research in Mathematics Education*, **18**, 141-157.
- BROWN, T. L., & CARR, T. H. (1989). Automaticity in skill acquisition: Mechanism for reducing interference in concurrent performance. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 686-700.
- CAMPBELL, J. I. D. (1987a). Network interference and mental multiplication. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 349-364.
- CAMPBELL, J. I. D. (1987b). Production, verification, and priming of multiplication facts. *Memory & Cognition*, **15**, 349-364.
- CAMPBELL, J. I. D., & GRAHAM, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Psychology*, **39**, 338-366.
- CANTOR, J., & ENGLE, R. W. (1993). Working-memory capacity as long-term memory activation: An individual-differences approach. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1101-1114.
- CARLSON, R. A., SULLIVAN, M. A., & SCHNEIDER, W. (1989). Practice and working memory effects in building procedural skill. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 517-526.
- CARPENTER, T. P., & MOSER, J. M. (1984). The acquisition of addition and subtraction concepts in grades one through three. *Journal for Research in Mathematics Education*, **15**, 179-202.
- CASE, R. (1985). *Intellectual development: Birth to adulthood*. New York: Academic Press.
- CASE, R., KURLAND, M., & GOLDBERG, J. (1982). Operational efficiency and the growth of short term memory. *Journal of Experimental Child Psychology*, **33**, 386-404.
- COMPTON, B. J., & LOGAN, G. D. (1991). The transition from algorithm to retrieval in memory-based theories of automaticity. *Memory & Cognition*, **19**, 151-158.
- CONWAY, A. R. A., & ENGLE, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, **123**, 354-373.
- ENGLE, R. W., CANTOR, J., & CARULLO, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 972-992.
- FAYOL, M., LARGY, P., & LEMAIRE, P. (1994). When cognitive load enhances subject-verb agreement errors: A study in French written language. *Quarterly Journal of Experimental Psychology*, **47A**, 437-464.
- FUSON, K. C. (1982). An analysis of the counting on solution procedure in addition. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 67-81). Hillsdale, NJ: Erlbaum.
- FUSON, K. C. (1984). More complexities in subtraction. *Journal for Research in Mathematics Education*, **15**, 212-225.
- FUSON, K. C., & KWON, Y. (1984). Korean children's single digit addition and subtraction: Numbers structured by tens. *Journal for Research in Mathematics Education*, **23**, 148-165.
- FUSON, K. C., RICHARDS, J., & BRIARS, D. J. (1982). The acquisition and elaboration of the number word sequence. In C. J. Brainerd (Ed.), *Children's logical and mathematical cognition: Progress in cognitive developmental research* (pp. 33-92). New York: Springer-Verlag.
- GEARY, D. C., & BURLINGHAM-DUBREE, M. (1989). External validation of the strategy choice model for addition. *Journal of Experimental Child Psychology*, **47**, 175-192.
- GEARY, D. C., FRENCH, P. A., & WILEY, J. G. (1993). Simple and complex mental subtraction: Strategy choice and speed-of-processing differences in young and elderly adults. *Psychology & Aging*, **8**, 242-256.
- GEARY, D. C., & WILEY, J. G. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in young & elderly adults. *Psychology & Aging*, **6**, 474-483.
- GROEN, G. J., & PARKMAN, J. M., (1972). A chronometric analysis of simple addition. *Psychological Review*, **79**, 329-343.
- GROEN, G. J., & RESNICK, L. B. (1977). Can preschool children invent addition algorithm? *Journal of Educational Psychology*, **69**, 645-652.

- HALFORD, G. S. (1993). *Children's understanding*. Hillsdale, NJ: Erlbaum.
- HAMANN, M. S., & ASHCRAFT, M. H. (1985). Simple and complex mental addition across development. *Journal of Experimental Child Psychology*, **40**, 49-70.
- HASHER, L., & ZACKS, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, **108**, 356-388.
- ILG, F., & AMES, L. B. (1951). Developmental trends in arithmetic. *Journal of Genetic Psychology*, **79**, 3-28.
- JUST, M. A., & CARPENTER, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, **99**, 122-149.
- KLAPP, S. T., BOCHES, C. A., TRABERT, M. L., & LOGAN, G. D. (1991). Automatizing alphabet arithmetic: II. Are there practice effects after automaticity is achieved? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 196-209.
- LARGY, P., FAYOL, M., & LEMAIRE, P. (1996). The homophone effect in written French: The case of verb-noun inflection errors. *Language & Cognitive Processes*, **11**, 217-255.
- LEFEVRE, J.-A., KULAK, A. G., & BISANZ, J. (1991). Individual differences and developmental change in the associative relations among numbers. *Journal of Experimental Child Psychology*, **52**, 256-274.
- LEFEVRE, J.-A., SADESKY, G. S., & BISANZ, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 216-230.
- LEMAIRE, P., ABDI, H., & FAYOL, M. (1996). The role of working memory resources in simple cognitive arithmetic. *European Journal of Cognitive Psychology*, **8**, 73-103.
- LOGAN, G. D. (1988a). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors*, **30**, 583-598.
- LOGAN, G. D. (1988b). Toward an instance theory of automatization. *Psychological Review*, **95**, 492-527.
- LOGAN, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 883-914.
- LOGAN, G. D., & KLAPP, S. T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 179-195.
- MILLER, K., PERLMUTTER, M., & KEATING, D. (1984). Cognitive arithmetic: Comparison of operations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 46-60.
- PASCUAL LEONE, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica*, **32**, 301-345.
- PASCUAL LEONE, J. (1988). Organismic processes for neo-Piagetian theories: A dialectical causal account of cognitive development. In A. Demetriou (Ed.), *The neo-Piagetian theories of cognitive development: Toward an integration* (pp. 25-64). Amsterdam: North-Holland.
- POSNER, M. I., & SNYDER, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55-85). Hillsdale, NJ: Erlbaum.
- SIEGLER, R. S. (1986). Unities across domains in children's strategy choices. In M. Perlmutter (Ed.), *Perspectives for intellectual development: The Minnesota symposia on child psychology* (Vol. 19, pp. 1-48). Hillsdale, NJ: Erlbaum.
- SIEGLER, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, **116**, 250-264.
- SIEGLER, R. S. (1989). Hazards of mental chronometry: An example from children's subtraction. *Journal of Educational Psychology*, **81**, 497-506.
- SIEGLER, R. S., & SHRAGER, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229-293). Hillsdale, NJ: Erlbaum.
- STEINBERG, R. M. (1985). Instruction on derived facts strategies in addition and subtraction. *Journal for Research in Mathematics Education*, **16**, 337-355.
- SVENSON, O. (1985). Memory retrieval of answers of simple additions as reflected in response latencies. *Acta Psychologica*, **59**, 285-304.
- SVENSON, O., & BROQUIST, S. (1975). Strategies for solving simple addition problems: A comparison of normal and subnormal children. *Scandinavian Journal of Psychology*, **16**, 143-151.
- SVENSON, O., & HEDENBORG, M. L. (1979). Strategies used by children when solving simple subtractions. *Acta Psychologica*, **43**, 477-489.
- WOOD, S. S., RESNICK, L. B., & GROEN, G. R. (1975). Experimental test of five process models for subtraction. *Journal of Educational Psychology*, **67**, 17-21.
- ZBRODOFF, N. J. (1995). Why is 9+7 harder than 2+3? Strength and interference as explanations of the problem-size effect. *Memory & Cognition*, **23**, 689-700.

NOTES

1. Of course, the data in the literature are primarily related to additions of numbers between 1 and 9 and it is possible that additions of larger numbers are not automatized, even in adult subjects. This is the reason why we have limited ourselves to small operands (from 1 to 4), for which the majority of adults probably employ a retrieval strategy.
2. This alternative hypothesis was suggested to us by E. Sieroff, to whom we are most grateful.

APPENDIX

Table A1
Summarized Table of ANOVAs

Sources	<i>df</i>	<i>MS_e</i>	<i>F</i>
Level	1,38	0.7506	77.54‡
Material	1,38	0.2307	119.80‡
Operation	1,38	0.1919	22.87‡
Operation span (OS)	3,114	0.1150	54.87‡
Item span (IS)	3,114	0.1984	186.44‡
Level × Material	1,38	0.2307	16.93‡
Level × Operation	1,38	0.1919	0.03
Level × OS	3,114	0.1150	8.23‡
Level × IS	3,114	0.1984	47.99‡
Material × Operation	1,38	0.2021	2.50
Level × Material × Operation	1,38	0.2021	3.41
Material × OS	3,114	0.1058	7.23‡
Level × Material × OS	3,114	0.1058	0.64
Material × IS	3,114	0.1525	6.25‡
Level × Material × IS	3,114	0.1525	7.81‡
Operation × OS	3,114	0.0945	2.03
Level × Operation × OS	3,114	0.0945	4.71†
Operation × IS	3,114	0.1005	2.20
Level × Operation × IS	3,114	0.1005	4.36†
OS × IS	9,342	0.1064	4.21‡
Level × OS × IS	9,342	0.1064	3.49‡
Material × Operation × OS	3,114	0.0732	0.98
Level × Material × Operation × OS	3,114	0.0732	0.89
Material × Operation × IS	3,114	0.1246	0.41
Level × Material × Operation × IS	3,114	0.1246	3.05*
Material × OS × IS	9,342	0.1038	1.60
Level × Material × OS × IS	9,342	0.1038	0.86
Operation × OS × IS	9,342	0.0923	1.52
Level × Operation × OS × IS	9,342	0.0923	1.06
Material × Operation × OS × IS	9,342	0.0867	0.85
Level × Material × Operation × OS × IS	9,342	0.0867	1.36

* $p < .05$. † $p < .01$. ‡ $p < .001$.

(Continued on next page)

Table A2
Tables of ANOVAs Performed for each Grade Separately

Sources	<i>df</i>	<i>MS_e</i>	<i>F</i>
Grade 2			
Material	1,19	0.219	24.55‡
Operation	1,19	0.215	9.47†
Operation span (OS)	3,57	0.097	62.41‡
Item span (IS)	3,57	0.212	197.94‡
Material × Operation	1,19	0.160	0.04
Material × OS	3,57	0.119	1.98
Material × IS	3,57	0.156	2.96*
Operation × OS	3,57	0.095	2.80*
Operation × IS	3,57	0.125	2.93*
OS × IS	9,171	0.113	5.44‡
Material × Operation × OS	3,57	0.085	0.62
Material × Operation × IS	3,57	0.176	0.87
Material × OS × IS	9,171	0.095	2.04*
Operation × OS × IS	9,171	0.079	1.67
Material × Operation × OS × IS	9,171	0.085	1.56
Grade 12			
Material	1,19	0.242	180.01‡
Operation	1,19	0.169	13.97†
OS	3,57	0.133	9.02†
IS	3,57	0.185	24.61‡
Material × Operation	1,19	0.244	4.87*
Material × OS	3,57	0.092	6.47‡
Material × IS	3,57	0.149	11.28‡
Operation × OS	3,57	0.094	3.94*
Operation × IS	3,57	0.076	3.86*
OS × IS	9,171	0.099	2.02*
Material × Operation × OS	3,57	0.073	1.38
Material × Operation × IS	3,57	0.073	3.82*
Material × OS × IS	9,171	0.113	0.55
Operation × OS × IS	9,171	0.105	1.00
Material × Operation × OS × IS	9,171	0.089	0.67

* $p < .05$. † $p < .01$. ‡ $p < .001$.

(Manuscript received July 22, 1996;
 revision accepted for publication January 11, 1997.)